

## AMENDMENTS TO THE SPECIFICATION

Please amend the paragraph beginning at line 14 on page 2 as follows:

As described above, automated monitoring of people and objects is useful in many applications such as security and surveillance. For example, automated monitoring of customers may be relevant to retail store managers who might wish to improve the layout of their stores through a better understanding of shopper behavior. Currently, due to the shortcomings of the current classes of object recognition methods, retail stores often use employees or consultants to monitor shopper activity rather than automated monitoring. Human monitoring also has shortcomings, such as human error and the cost of employing additional personnel ~~personal~~. Furthermore, in security applications it is typically necessary ~~[[to]]~~ for automated monitoring to provide highly accurate and prompt analysis to provide maximum safety. However, due to the limitations of current automated monitoring methods, accuracy and/or prompt response time may not be provided, reducing the effectiveness and safety provided by current automated monitoring methods.

Please amend the title at line 1 on page 6 as follows:

DETAILED DESCRIPTION ~~BEST MODE FOR CARRYING OUT THE~~  
INVENTION

Please amend the paragraph beginning at line 5 on page 15 as follows:

Embodiments of the invention also allow for further discretization of space along the third, Z-dimension, as shown in Figure 5. Figure 5 illustrates a diagrammatic perspective view 500 of the three-dimensional coordinate system of Figure 4 with the three-dimensional point cloud 330 and vertical bins 430

being divided horizontally into multiple slices in accordance with an embodiment of the present invention. Within 3D world coordinate system 440, each vertical bin 430 may be divided into several box-shaped sub-bins, by introducing dividing planes that are parallel to the ground-level plane. This creates two or more slices 510 of boxes, where the boxes in a particular slice 510 all have centers lying in some plane parallel to the ground-level plane. Plan-view images of any type can then be constructed, including all of the types discussed above, for one or more of these slices 510. Each plan-view image of a particular slice has one pixel per box in that slice image. One slice of boxes is highlighted in Figure 5.

Please amend the paragraph beginning at line 11 on page 17 as follows:

With reference to Figure 1C, raw plan-view templates 175 are modified by template processor 180 to produce plan-view templates 125. It should be appreciated that template processor 180 is optional, and is not required for performing plan-view template generation 128. Template processor 180 may ~~480~~ process raw plan-view templates 175 to produce new data representations embodied by plan-view templates 125. This processing may ~~[[be]]~~ include a combination of one or more of many types of data normalization and transformation, including but not limited to scaling in the spatial dimensions of the raw plan-view template, rotating the raw plan-view template image data, removal of small isolated regions of non-zero raw plan-view template data, smoothing of the raw plan-view template data, convolution of the raw plan-view template data with an image kernel, interpolation of the raw plan-view template data across small regions of zero or unreliable data, and representation of the raw plan-view template data in terms of contours, spatial moments, basis functions or vectors, or other primitives.

Please amend the paragraph beginning at line 5 on page 20 as follows:

With reference to Figure 1A, plan-view templates 125 are then processed by one or more classifiers 130 to make decisions 135 about the type, identity, and/or pose of the object(s) corresponding to the data in the plan-view templates. Classifiers 130 may be binary (e.g., provide a “yes” or “no” answer), or multi-class (e.g., assigning the input to one of K classes, where K may be greater than 2). Each classifier decision 135 may be based upon input comprising one or more plan-view templates 125. When multiple templates are provided to a single classifier to make a single decision, these templates may be extracted from different portions of the same plan-view image, from different types of plan-view images created from depth data 110 and optional non-depth data 115 pertaining to a single time instance, from plan-view images created from different horizontal slices of the same three-dimensional point cloud of data, from plan-view images created from depth data 110 and optional non-depth data 115 obtained at different time instances, or some combination thereof.

Please amend the paragraphs beginning at line 6 on page 25 as follows:

At step 720, calibration information of the visual sensor is received. In one embodiment, the calibration information includes vertical and horizontal visual sensor focal lengths, image pixel coordinates of the visual sensor's center of projection, the location and orientation of different imaging elements of the visual sensor relative to that of a reference imaging element, and an indication of the visual sensor's location and orientation. At step 730, a plan-view image is generated based on the depth data. In one embodiment, the plan-view image is generated according to a process described with reference to 730 of Figure 8.

Figure 8 illustrates a flow chart of ~~a process~~ expanding on step 730 of Figure 7 for generating a plan-view images in accordance with an embodiment of the present invention. In one embodiment, ~~step 730 process 800~~ is implemented by a plan-view image generator (e.g., plan-view image generator 120 of Figures 1A and 1B). At step 810, a subset of points of the image is selected. In one embodiment, the subset of points is selected according to foreground segmentation.

Please amend the paragraphs beginning at line 7 on page 26 as follows:

At step 840, at least a portion of the three-dimensional point cloud is mapped into at least one plan-view image based on the points' three-dimensional coordinates and optional associated non-depth data. The plan-view images provide two-dimensional representations of the three-dimensional point cloud. In one embodiment, the portion comprises at least one horizontal slice of the three-dimensional point cloud. In another embodiment, the portion comprises the entire three-dimensional point cloud. In one embodiment, a portion of the plan-view image is mapped according to a process described with reference to 840 of Figure 9.

Figure 9 illustrates a flow chart of ~~a process~~ expanding on step 840 of Figure 8 for mapping at least a portion of a three-dimensional point cloud in accordance with an embodiment of the present invention. In one embodiment, ~~step 840 process 900~~ is implemented by a plan-view projector (e.g., plan-view projector 160 of Figure 1B). At step 910, the three-dimensional point space is partitioned into a plurality of vertically oriented bins. At step 920, one or more statistics of the points within each vertical bin are computed. If step 830 of Figure 8 has been used to divide the three-dimensional space into horizontal

slices, statistics may be computed separately for the portions of vertical bins lying in different horizontal slices. At step 930, each computed statistic of the points in the vertically oriented bins, optionally restricted to one or more horizontal slices, are mapped into corresponding plan-view images, wherein each pixel of a plan-view image corresponds to one of the vertical bins.